EDITORIALS

Validating and Improving Models for Projecting the Absolute Risk of Breast Cancer

Mitchell H. Gail, Joseph P. Costantino

In this issue of the Journal, Rockhill et al. *(1)* use data from 1992 to 1997 from white women in the Nurses' Health Study (NHS) to address the validity and uses of the model of Gail et al. *(2)* for estimating breast cancer risk. One issue is how well the model predicts the risk of invasive breast cancer in various subgroups of women ("calibration"). A second issue is how sharply the model discriminates women who will develop breast cancer from those who will not ("discriminatory accuracy"). Rockhill et al. also comment on applications of the model.

The model of Gail et al. *(2)* was based on follow-up of white women in the Breast Cancer Detection Demonstration Project (BCDDP). A nested case–control study in the BCDDP was used to estimate multivariate relative risks based on age at menarche, age at first live birth, number of first-degree relatives (mother and sisters) with breast cancer, number of breast biopsies, and whether or not atypical hyperplasia was present on any biopsy specimen. To estimate the absolute risk of developing breast cancer over a specified age interval, data from the case–control study were combined with age-specific breast cancer rates in the entire BCDDP cohort. Because BCDDP participants were screened annually with mammography, it was anticipated *(2,3)* and found *(4)* that the model would overestimate risk in unscreened younger women, especially in the era before 1982 when screening was not available routinely *(5)*.

In preparation for the Breast Cancer Prevention Trial (BCPT) *(6),* C. K. Redmond and S. Anderson (University of Pittsburgh, PA) replaced BCDDP age-specific rates with rates from 1983 through 1987 for invasive breast cancer from the Surveillance, Epidemiology, and End Results[1] Program of the National Cancer Institute (NCI). They also provided separate estimates for white and black women *(7)*. This modified model is available as the "NCI Risk Disk" from NCI's Office of Communications and can be accessed through http://cancernet.nci.nih.gov/genetics_prevention.html. This is the version of the Gail et al. model studied by Rockhill et al.

Previous validation efforts *(7)* showed that the relative risk features of the model have been consistently replicated across several studies, including the Cancer and Steroid Hormone Study, BCPT, and an earlier NHS report *(5)*. This consistency now extends to the NHS data from 1992 to 1997 presented by Rockhill et al.

The 1354 incident breast cancers studied by Rockhill et al. add appreciably to what is known about how well the Gail et al. model predicts risk in white women with access to screening mammography. Rockhill et al. demonstrated good agreement between observed (O) and expected (E) cancers. The overall E/O ratio was 0.94 (95% confidence interval [CI] = 0.89 to 0.99). In women with a predicted 5-year risk of less than 1.67%, the E/O ratio was 0.86 (95% CI = 0.80 to 0.92); in those with a higher predicted risk, the E/O ratio was 1.04 (95% CI = 0.96 to 1.12). These data reinforce the findings of Costantino et al. *(7),* who found that the E/O ratio was 1.03 (95% CI = 0.88 to 1.21) overall.

Thus, available data indicate that the modified model of Gail et al. predicts risk well (good calibration) for white women with access to mammography. Nonetheless, the model does not take certain types of information into account. For example, the model underestimates risk in women with a history of breast cancer, lobular carcinoma *in situ,* or ductal carcinoma *in situ* and should not be used for such women. Recent immigrants from rural China or Japan probably have lower risk than predicted by the model [*see* references in *(8)*]. The model may underestimate risk for women with demonstrated mutations of the BRCA1 or BRCA2 genes *(9)* or women with Cowden syndrome *(10)* or the Li–Fraumeni syndrome *(11)*. A major uncertainty concerns the performance of the model in African-American women, Hispanic women, and other subgroups for whom there are scant validation data. The model may underestimate risk in African-American women *(12)*. For these reasons, we recommend that a knowledgeable counselor assist in interpreting results from the model.

Rockhill et al. present data indicating that, even if the model were well calibrated, it had only modest discriminatory accuracy to predict which individual women will develop breast cancer and which will not. In their data, the women classified in the highest decile of risk were only 2.8 times as likely to develop disease as those in the lowest decile of risk. A highly discriminating model would assign very high risks precisely to those women who will develop disease and very low risk to those women who will not develop the disease. The more discriminating the model is, the more effectively intervention strategies can be targeted to individual women.

Well-calibrated projections of absolute risk, even from models with modest discriminatory accuracy, have several important uses. They are used to plan intervention trials, such as the BCPT, because the power of such studies depends on the numbers of incident breast cancers, a reflection of average absolute risk. The model has been used to counsel and educate women, who often overestimate their risk of breast cancer, and to assist in making clinical decisions. For example, if a woman has a well-calibrated 5-year risk of breast cancer of 1.0% and of stroke of 2.0% and if she is told that tamoxifen can lower her risk of breast cancer to 0.5%, while raising her risk of stroke to 3.2%, she should

Affiliations of authors: M. H. Gail, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD; J. P. Costantino, Graduate School of Public Health, University of Pittsburgh, PA.

Correspondence to: Mitchell H. Gail, M.D., Ph.D., National Institutes of Health, EPS-8032, Bethesda, MD 20892 (gailm@exchange.nih.gov).

See "Note" following "References."

think carefully about the pros and cons of taking tamoxifen *(12).* Absolute risk projections might also convince a woman in her 40s to begin mammographic screening if she had the same risk of breast cancer as a 50-year-old woman *(13).*

Rockhill et al. assess the public health impact of the preventive use of tamoxifen. Only a small proportion of NHS participants stand to benefit from tamoxifen, if one weighs the risk and benefits as suggested by Gail et al. *(12).* Nonetheless, these small proportions could represent large absolute numbers of women who benefit. Moreover, one hopes to find interventions with fewer toxic effects, for which larger proportions of women could benefit. Efforts to weigh risks and benefits will require well-calibrated models not only for breast cancer but also for other health outcomes affected by the intervention, such as stroke.

Rockhill et al. focus on 5-year risk projections. Clinical decisions may depend on longer time frames. For this reason, it is often useful to present projections over various intervals. A risk that seems small over a 5-year period can compound over time. For example, a 40-year-old white nulliparous woman whose mother had breast cancer, who began menstruating at age 12 years, and who had one breast biopsy specimen with atypical hyperplasia would have projected risks of 3.4%, 8.1%, 16.5%, and 26.6%, respectively, over periods of 5, 10, 20, and 30 years. It would be useful to know how well the model is calibrated and what its discriminatory accuracy is over longer time intervals.

Thus, Rockhill et al. provide important evidence that the modified model of Gail et al. is well calibrated for white women, and they challenge researchers to develop more discriminating models. As they point out, there may be opportunities to improve the model by using more powerful risk predictors. Some promising predictors include mammographic density *(14,15),* more detailed data on family history of breast or ovarian cancer *(16,17)* or reproductive history *(18),* data on levels of plasma estrogen *(19),* cytology from breast duct lavage, and genetic or other molecular markers. It will take years to determine the utility and validity of models based on such predictors, however. In the meantime, the current model has its uses. It is likely that a simple model based on data easily acquired from the medical history, such as the model of Gail et al., will remain useful for identifying women who could benefit from preventive interventions or who require more discriminating but more invasive or costly tests to predict risk.

REFERENCES

(1) Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. J Natl Cancer Inst 2001;93:358–66.

(2) Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 1989; 81:1879–86.

(3) Gail MH, Benichou J**.** Validation studies on a model for breast cancer risk. J Natl Cancer Inst 1994;86:573–5.

(4) Bondy ML, Lustbader ED, Halabi S, Ross E, Vogel V. Validation of a breast cancer risk assessment model in women with a positive family history. J Natl Cancer Inst 1994;86:620–5.

(5) Spiegelman D, Colditz GA, Hunter D, Hertzmark E. Validation of the Gail et al. model for predicting individual breast cancer risk. J Natl Cancer Inst 1994;86:600–7.

(6) Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, et al. Tamoxifen for prevention of breast cancer: report of the National Breast and Bowel Project P-1 study. J Natl Cancer Inst 1998;90: 1371–88.

(7) Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. J Natl Cancer Inst 1999;91:1541–8.

(8) Gail MH, Benichou J. Assessing the risk of breast cancer in individuals. In: DeVita VT Jr, Hellman S, Rosenberg SA, editors. Cancer prevention. Philadelphia (PA): Lippincott; 1992. p. 1–15.

(9) Struewing JP, Hartge P, Wacholder S, Baker SM, Berlin M, McAdams M, et al. The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. N Engl J Med 1997;336:1401–8.

(10) Hanssen AM, Fryns JP. Cowden syndrome. J Med Genet 1995;32:117–9.

(11) Li F. Familial aggregation. In: Schottenfeld D, Fraumeni JF Jr, editors. Cancer epidemiology and prevention. 2nd ed. New York (NY): Oxford University Press; 1996. p. 546–58.

(12) Gail MH, Costantino JP, Bryant J, Croyle R, Freedman L, Helzlsouer K, et al. Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. J Natl Cancer Inst 1999;91:1829–46.

(13) Gail M, Rimer B. Risk-based recommendations for mammographic screening for women in their forties. J Clin Oncol 1998:16;3105–14.

(14) Boyd NF, Byng JW, Jong RA, Fishell EK, Little LE, Miller AB, et al. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. J Natl Cancer Inst 1995;87:670–5.

(15) Byrne C, Schairer C, Wolfe J, Parekh N, Salane M, Brinton LA, et al. Mammographic features and breast cancer risk: effects with time, age, and menopause status. J Natl Cancer Inst 1995;87:1622–9.

(16) Claus EB, Risch N, Thompson WD. Autosomal dominant inheritance of early-onset breast cancer: Implications for risk prediction. Cancer 1994;73: 643–51.

(17) Berry DA, Parmigiani G, Sanchez J, Schildkraut J, Winer E. Probability of carrying a mutation of breast–ovarian cancer gene BRCA1 based on family history. J Natl Cancer Inst 1997;89:227–38.

(18) Rosner B, Colditz GA. Nurses' Health Study: log-incidence mathematical model of breast cancer incidence. J Natl Cancer Inst 1996;88:359–64.

(19) Hankinson SE, Willett WC, Manson JE, Colditz GA, Hunter DJ, Spiegelman D, et al. Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. J Natl Cancer Inst 1998;90:1292–9.

NOTE

[1]*Editor's note:* SEER is a set of geographically defined, population-based, central cancer registries in the United States, operated by local nonprofit organizations under contract to the National Cancer Institute (NCI). Registry data are submitted electronically without personal identifiers to the NCI on a biannual basis, and the NCI makes the data available to the public for scientific research.